Why Data Linkage?

PB98-163769

JUNE 1996

THE IMPORTANCE OF CODES
(CRASH OUTCOME DATA EVALUATION SYSTEM)

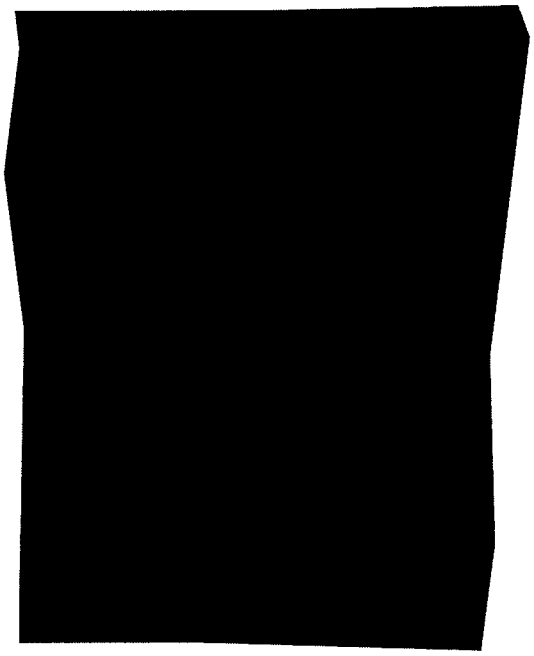
Why Data Linkage?

THE IMPORTANCE OF CODES
(CRASH OUTCOME DATA EVALUATION SYSTEM)
JUNE 1996

Table of Contents

Introduction

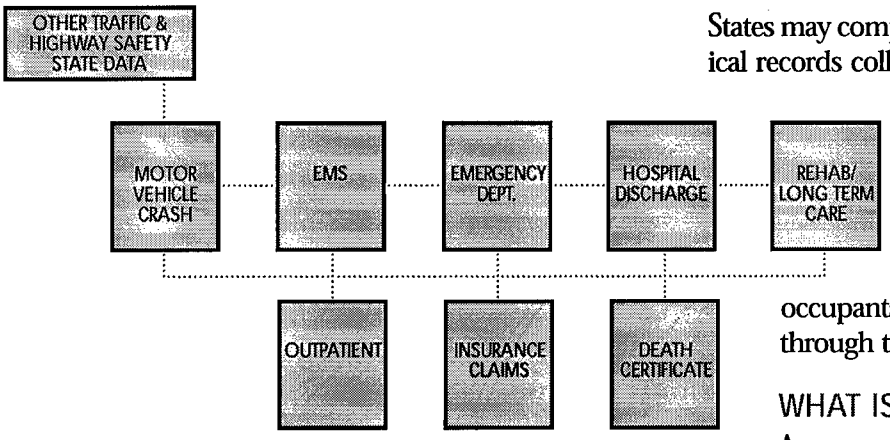
Injuries resulting from motor vehicle crashes remain a major public health problem. Over five million people were injured and more than 41,000 killed in 1995. Injuries cost society annually $151 billion, of which $17 billion are medical expenses. Seventy percent of these costs are paid by third party payers. These costs cause an unnecessary burden of increased taxes and insurance premiums. Efforts to control injury must expand to reflect the magnitude of this challenge.



Motor vehicle injuries can be prevented but only if we understand fully the nature of the problem. What environmental factors are associated with injuries? What can be done to reduce injury severity? Which injuries lead to long term disabilities? Which injuries are associated with high health care costs? If we know what causes poor medical outcome and high health care costs, then efforts to reduce these problems can be more effectively designed and evaluated.

Data linkage is the key to providing the information needed to understand what we should focus on first.

EXHIBIT 1



WHAT IS DATA LINKAGE?

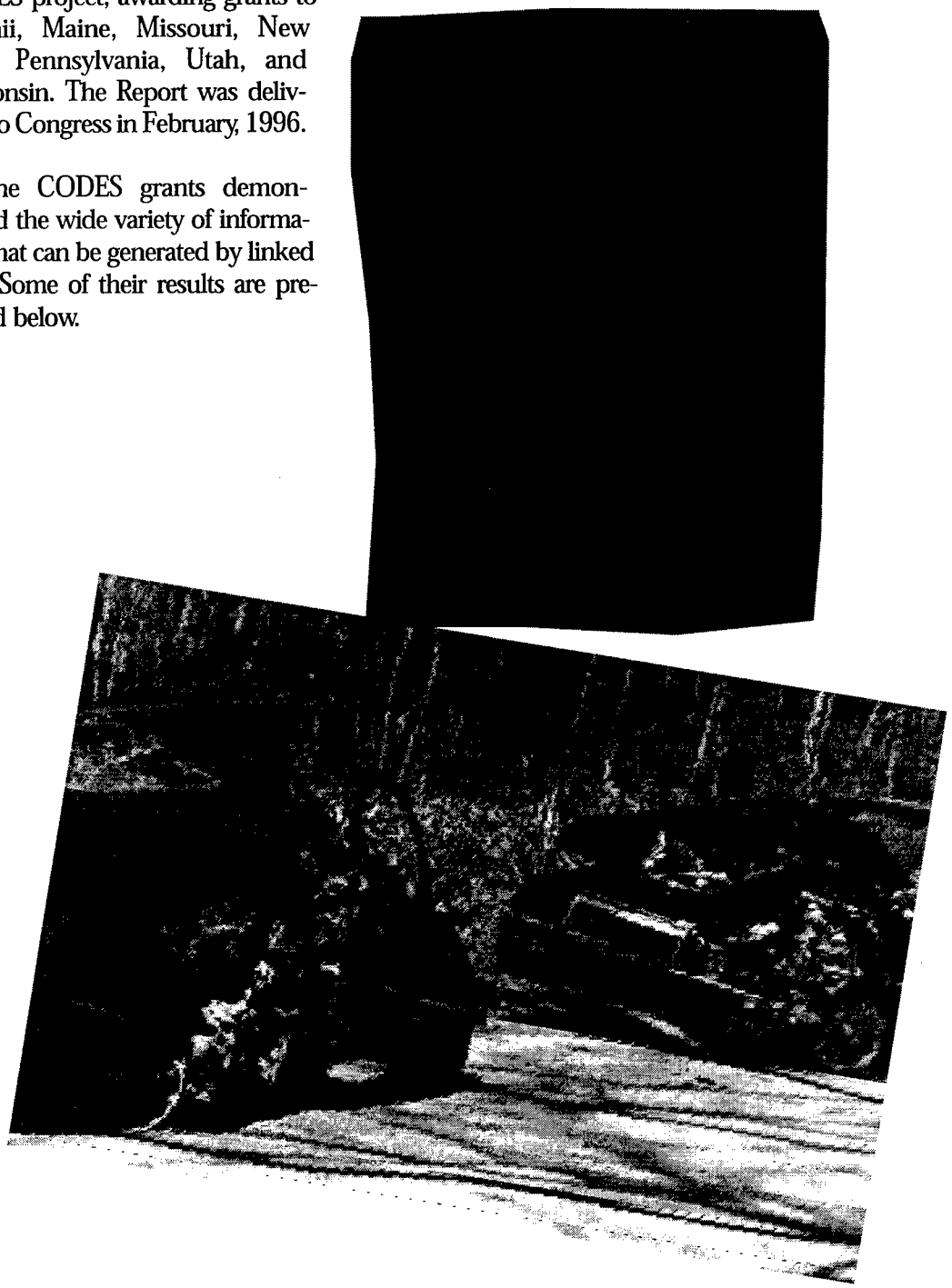
States may computerize statewide crash and medical records collected at the crash scene, enroute, at the emergency department, in the hospital, and after discharge for occupants involved in a motor vehicle crash. When these different records are linked, as in Exhibit 1, injured occupants can be traced from the scene through the health care system.

WHAT IS CODES?

An example of linked data is the Crash Outcome Data Evaluation System project

which evolved from the Intermodal Surface Transportation Efficiency Act of 1991 (ISTEA). ISTEA mandated that the National Traffic Safety Administration (NHTSA) prepare a Report to Congress about the benefits of safety belt and motorcycle helmet use. To obtain the crash and injury outcome information needed for this report, NHTSA sponsored the CODES project, awarding grants to Hawaii, Maine, Missouri, New York, Pennsylvania, Utah, and Wisconsin. The Report was delivered to Congress in February, 1996.

The CODES grants demonstrated the wide variety of information that can be generated by linked data. Some of their results are presented below.



What do Linked Crash and Injury Data Tell Us About Who is at Risk for Increased Severity or High Health Care Costs?

Crash data alone are unable to convey the injury problem in terms of the magnitude of the medical and financial consequences. By linking the crash, vehicle, and occupant behavior characteristics to their specific medical and financial outcomes we can identify the factors which increase risk.

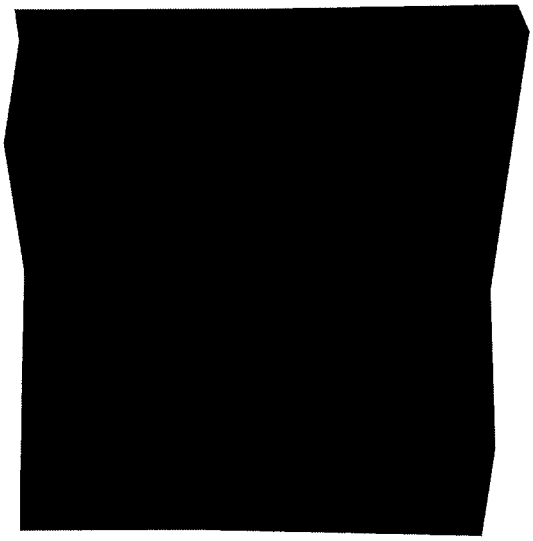
• POPULATIONS AT RISK

Linked data are useful for targeting the populations at risk for the different types of injuries attributed to motor vehicle crashes.

- *Wisconsin reported that some age groups—the young and the older drivers—are at greater risk for high costs when crashes occur. The young are over- represented among the pedestrians, bicyclists, and motorcyclists. Older drivers are more likely to be injured inside a vehicle and to experience chest injuries, which have higher costs.*

- *Utah identified the costs to the state economy incurred by young, inexperienced drivers and the causes of these costs. Armed with real Utah-specific data, the 1997 Utah Legislature will be able to make an educated vote on the graduated driver licensing legislation it is expected to consider.*

- *Utah discovered that the rural areas with high tourism were burdened with increased health care costs because of the extra demand for service. In one county, tourists were responsible for 11.5 percent of all EMS transports, 18.5 percent of all rural inpatient charges, 14.5 percent of all rural outpatient charges and 69 percent of the fatalities.*



• HIGH COST INJURIES

Linked data are useful to identify the most expensive injuries resulting from motor vehicle crashes. Health care costs are more likely to be decreased when the high cost injuries are prevented.

- *Missouri observed that regardless of the specific crash characteristics, costs were highest for upper leg injuries and only slightly lower for pelvic, head/neck, abdomen, and chest injuries.*

- *New York evaluated liver and spleen injuries by crash severity, safety belt manufacturer, make, and model of car and provided the information to a safety belt manufacturer interested in improving product safety.*

Another group in New York is studying the types of injuries occurring to rear seat occupants and those riding in pick-up truck beds. This information will be used to inform the public and possibly support legislation against riding in a pick-up truck bed.

- *Maine studied injuries and their costs resulting from collisions with moose and deer, an increasing problem in the State. This information is being used to heighten public awareness of the dangers, especially during times of the year when moose and deer are more likely to be encountered.*

- *Maine also is linking ten years of EMS, hospital, and death certificate data to monitor the occurrence of head and spinal cord injuries, and to evaluate the outcomes of specific treatment protocols designed to reduce disability from this type of injury. The results will be used to improve the quality of medical care.*

• HIGH COST CRASH TYPES AND LOCATIONS

Linked data are useful to identify the types and locations of motor vehicle crashes that are likely to cause the most expensive injuries.

- *Hawaii developed a structural model of the relationships between crash type and injury levels. The modeling has focused on collisions involving bicycles, mopeds, pedestrians, older drivers, and tourists.*

- *Hawaii also developed a traffic safety geographic information system to map collisions and their characteristics. Current applications involve mapping the locations of collisions with utility poles and identifying the health and financial outcomes of these events.*

What Do Linked Data Tell Us About the Impact of Different Behaviors on Outcome?

• USING SAFETY BELTS

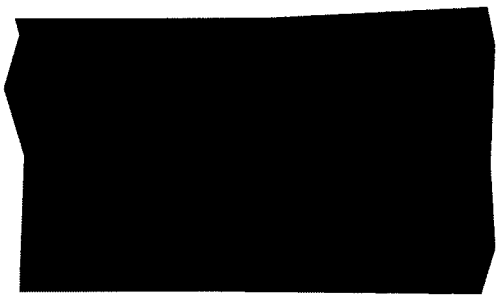
Linked data from the CODES project demonstrated that safety belts are 20 percent effective in preventing any type of injury, 60 percent effective in preventing death, and cause a downward shift in the severity of injury. CODES also reported that drivers who are unbelted, on average, cost an additional $5,000 in hospital inpatient charges. These results were derived by statistically merging the effectiveness rates from the seven CODES states into one effectiveness rate. The CODES states used their respective effectiveness rates to support belt use.

- *The Coalition for Utah Traffic Safety (CUTS) prepared the videotape "Stop Rolling The Dice" with CODES data. It was distributed to members of the Utah Legislature when a primary safety belt law was presented. The same video is being used in Utah automobile dealership showrooms as a point of sale promotion for safety belt use.*

- *A rural hospital in Utah compared the safety belt use and injury rates for its local community and then conducted a campaign to increase belt use to alleviate the trauma cases that heavily drain rural EMS and hospital resources.*

- *Missouri identified higher ejection rates, rates of serious injuries, and death rates for pickup truck drivers who have a lower belt use rate, 55 percent compared to 82 percent for passenger car drivers.*

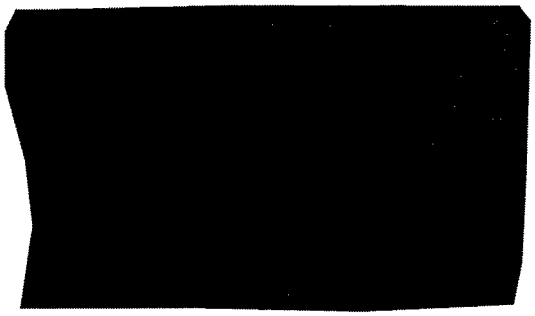
- *Missouri supported legislation mandating safety belt use in pickup trucks and demonstrated to business groups the cost burden associated with the failure to buckle up.*



• USING A MOTORCYCLE HELMET

Prior to having access to the medical diagnosis information generated by data

linkage, helmets were shown to be 34 percent effective in preventing death.



But CODES linked data showed that helmets are 67 percent effective in preventing brain injury and that about $15,000 of the average inpatient charge is saved during the first 12 months for every motorcycle rider who, because of helmet use, does not sustain a brain injury.

- *Missouri used its CODES data to prevent the weakening of Missouri's helmet law and to demonstrate to local groups how much the unhelmeted riders involved in crashes cost the State.*

• Driving Under the Influence

Alcohol is a factor in motor vehicle crashes. Linked data enable the risk to be quantified to indicate the percentage of health care costs caused by the alcohol-impaired driver.

- *Utah documented that during 1991, alcohol-related crashes cost a total of $5.6M in inpatient charges. Drivers younger than age 30 were responsible for 56 percent of the total inpatient charges resulting from alcohol involved crashes, but represented only 25 percent of the total drivers involved. Utah used these and other CODES data to document the positive impact and dollar savings that the State's DUI Diversion team had on targeted counties.*

• Running Red Lights

Linked data are useful to identify the types of injuries and their costs resulting from crashes caused by running red lights.

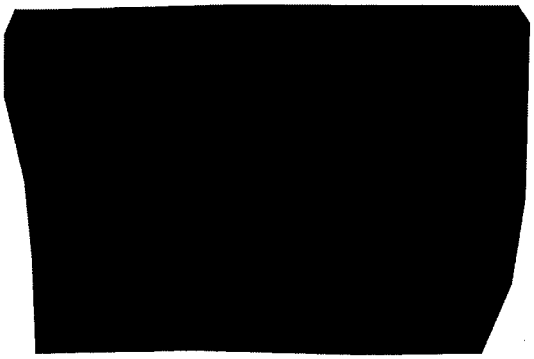
- *Utah's Safety Management System (SMS) identified a problem with "red light running" and, as a result, obtained funding from NHTSA to conduct a statewide campaign to reduce this problem.*

What Do Linked Data Tell Us About the Emergency Medical Services System?

Linkage of crash and injury data provides additional information about what happens before EMS arrives at the scene and after EMS delivers the patient to the next level of care. This information is useful for determining areas of increased risk.

• Where to Place Resources

Linked data are useful to determine the most effective placement of EMS resources to match the needs of the emergency victim.



- *Hawaii incorporated the location of EMS pick-ups into a geographic information system. Location was recorded as the nearest intersection/mile marker to the EMS pick-up. Population size, median income, percentage of elderly, age/gender, medical problems, and disposition were added to each location. EMS used this information to study injury patterns related to location.*

• The Impact of a Delay to EMS

Linkage provides EMS with time information to evaluate the impact of the delay to EMS.

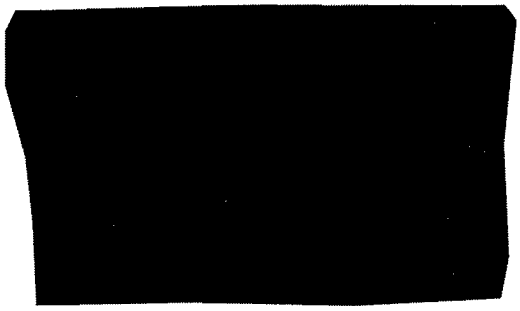
- *Missouri analyzed the distribution of injuries by regional planning areas. The linked data indicated that, in two rural areas, more than 50 percent of the patients suffered delays of more than 120 minutes before arriving at the hospital. These same patients were more likely to suffer poor outcomes, as defined by increased hospitalizations, higher severity, etc.*

- *New York found in one county that the risk for hospitalization and high costs increased 10 percent when the EMS response time was greater than 60 minutes. New York is using its CODES linked data to support development of the Automated Collision Notification (ACN) device which will automatically call 911 and report a motor vehi-*

cle collision, request emergency services and document the exact time and location of the collision. The activation of this device will be based upon the velocity of the vehicle, similar to the activation of an air bag.

- THE RELIABILITY OF THE EMERGENCY MEDICAL TECHNICIAN (EMT) ASSESSMENTS

The Emergency Medical Technician assesses the patient's symptoms and then determines the appropriate treatment. Linked data enable the EMT assessment to be compared to the final medical diagnoses determined by hospital personnel. This information is important to discriminate between when EMTs are making patient assessment or documentation errors and to target training accordingly.



- *Missouri found that the EMT assessment and hospital discharge diagnoses agreed for 64 percent of the patients. EMTs failed to document 29 percent of head injuries and 42 percent of the chest injuries requiring hospital admission. This information was used by EMS to improve the quality of EMS documentation.*

- THE EMS ROLE IN PUBLIC HEALTH

Linked data are useful to highlight the role of out-of-hospital care in the emergency medical care system and provide EMS at all levels with injury prevention information to raise public awareness about the benefits of safety belts and helmets.

- *Utah classified severity for persons injured in motor vehicle crashes and transported by EMS according to four levels of out-of-hospital treatment and then used it to measure the impact of safety belts on the need for EMS services. Safety belt users were less likely to need EMS treatment; for the belted who did receive EMS treatment, they were less likely to need bleeding control, cervical immobilization, intravenous placement, oxygen therapy, spinal immobilization, and ventilation than someone who was unbelted. Utah EMS is using these results to provide public education about the benefits of safety belts in reducing injury severity.*

What Do Linked Data Tell Us About Safety at the Community Level?

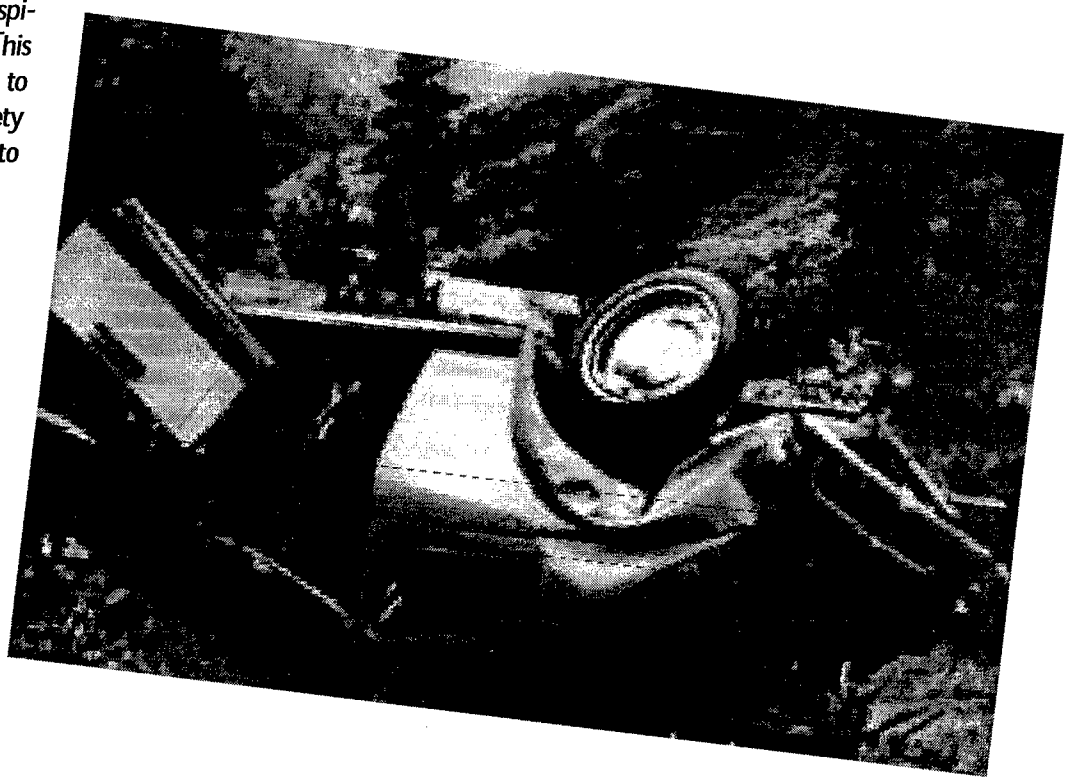
State data are really local data collected by police, EMS and other health care providers, businesses, insurers, and government that are merged to create the statewide data base. These data are population-based when they include everyone involved. They can be disaggregated from the state level back to the local level to generate information about which types of injuries and their causes increase the risk for severity and high costs at the local level. The disaggregated state data provide a context within which to compare the local area to similar areas or to the state as a whole. They also make it possible to identify local priorities based on local data.

- *Wisconsin integrated its CODES data as part of a State initiative to build safe communities. As part of the assessment of a community's traffic safety, CODES linked data were used to prepare a "Community Crash Cost Report." This Report indicated the number and cost of crash-related injury hospitalizations for residents of the community. A "Report Card" was also prepared indicating the community's experience over five years compared to the entire state.*

- *New York reported that in one county crashes were more likely to involve passenger cars, alcohol, failure to yield, head-on collision, multiple vehicles, and less likely to involve pedestrians or bicyclists, compared to the state as a whole. In addition, the occupants of the vehicles involved in these crashes were more likely to be injured, experience inpatient costs higher than $2,000, and pay higher insurance costs. Because the linkage included crash records for all occupants, regardless of whether an injury was documented, linked data also made it possible to discover for this county that the crash reports that linked to an injury record were less likely to document the injury on the crash report compared to the state results.*

- *New York also evaluated injury patterns by type of roadway (expressway, state route). An analysis of one specific town indicated that many more injuries occurred on the State road compared to the expressway. Within the county as a whole, 30 injuries were expected on the state roads and 44 actually occurred. New York reported this roadway specific information to a statewide planning committee, the Safety Management System, for use in setting priorities for improving traffic safety.*

- *Missouri found that groups of rural counties had lower belt use rates and higher risk for hospitalization and death. This information was used to justify increasing safety belt use programs to these areas.*



What Improvement in Data Quality Results From CODES?

Linkage highlights problems with data quality. Not all injuries are documented on the crash report. Some are not required to be reported because they occur as the result of crashes on private roads. Some are missed because of non-compliance with reporting requirements. Others are not documented because of delayed symptoms. Many of the missing injuries, although minor in terms of survival, may cause high health care costs. Linkage makes it possible to determine the significance of the missing information.

- *Utah reported that 10 percent of the occupants designated as not injured on the crash reports linked to at least one injury record. The addition of these cases was significant since their average cost and length of stay were as high as those for the occupants designated as severely injured.*

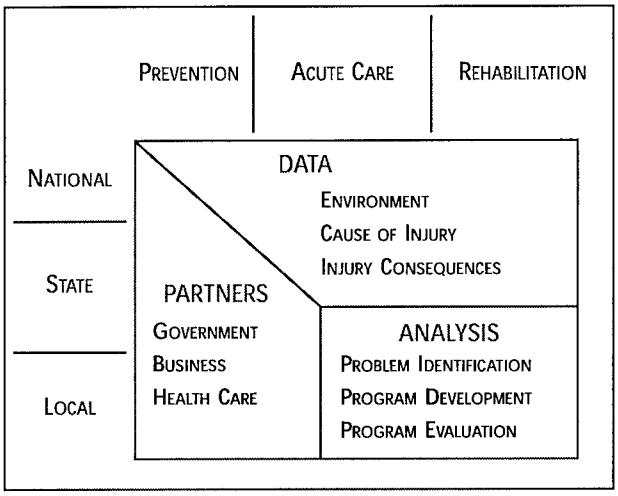
- *Hawaii reported that about 25 percent of the EMS and hospital discharge records and 45 percent of the claims records which documented an injured motor vehicle crash occupant failed to link to a crash report. Missing records were associated with specific geographical locations, time and date, age, severity and type of injury, and problems with the data processing. These data were shared with EMS.*

- *Maine compared the police, EMS, and hospital documentation of injury and evaluated the significance of the over-reporting of safety belt use. Recommendations were then made to improve the quality of the data used for linkage and analysis.*

What Benefits Result From the Collaboration Required to Implement CODES?

An injury control system depends on working partnerships between business, government and the health care community at the national, state, and local levels. Exhibit 2 shows that each has unique, yet complementary, responsibilities across the injury control system components of prevention, acute care, and rehabilitation. Starting at the local level and building through the state to the national levels, people, data collection and analysis are major parts of the fabric that bind these diverse activities together.

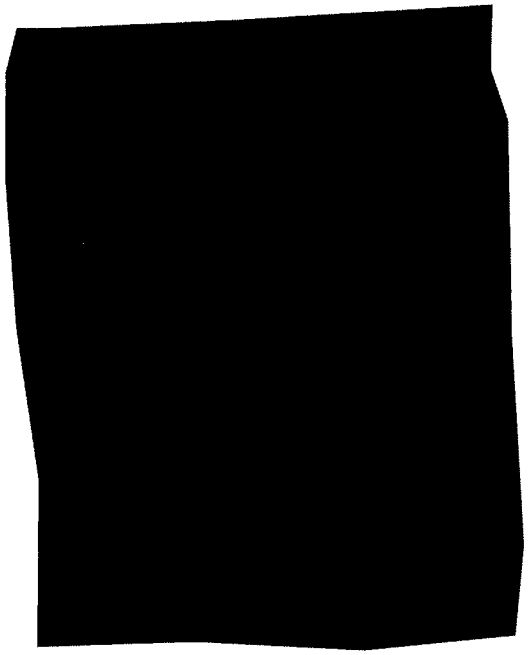
EXHIBIT 2



An Advisory Committee is recommended to serve as the mechanism to bring these components together for the purpose of facilitating data access, linkage, and analysis. Exhibit 3 indicates the types of representation recommended for such an Advisory Committee.

EXHIBIT 3

TYPES OF REPRESENTATION ON CODES ADVISORY COMMITTEE	
TRANSPORTATION	PHYSICIANS, NURSES, ETC.
MOTOR VEHICLES	EMS
TRAFFIC SAFETY	HOSPITAL
LAW ENFORCEMENT	REHABILITATION
ENGINEERS	LONG TERM CARE
VEHICLE INSURANCE	HEALTH AND HUMAN SERVICES
CONSUMER GROUPS	PUBLIC HEALTH/INJURY CONTROL
LEGISLATOR	HEALTH INSURANCE



What Are the Benefits of Probabilistic Linkage Techniques to Link CODES Data?

In CODES, crash and injury records are linked using probabilistic linkage techniques. These techniques enable valid pairs to be identified without requiring exact matches on the identifiers. As a result, large volumes of computerized state data may be successfully linked in a phenomenally short amount of time at relatively low cost.

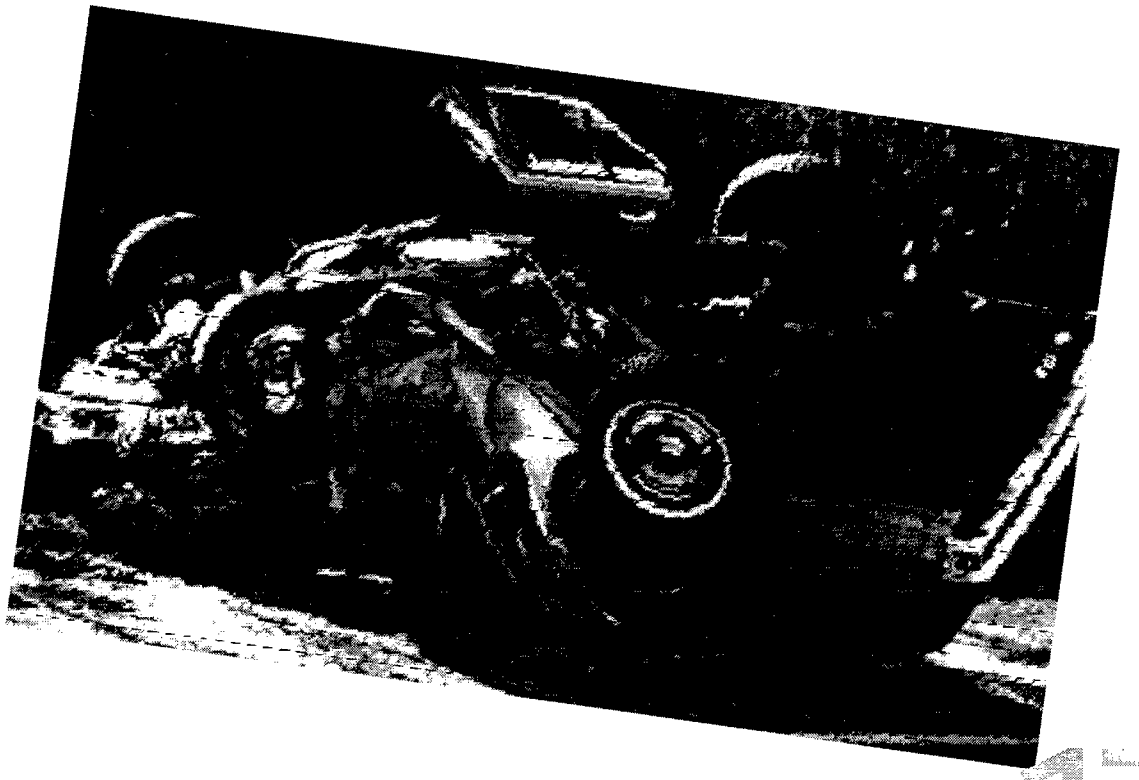
In this type of linkage, combinations of identifiers are used to identify records for specific individuals. Geographic location, times, type of vehicle and other variables are used to identify a specific crash. Age/date of birth, gender, description of injury, name or initials (if available) and other variables are used to locate a specific person. Names, when available, are converted to phonetic spellings to permit linkage. The patient identifiers are sanitized by the State prior to analysis and distribution. For more information on the probabilistic linkage techniques, see the CODES Technical Report, published by NHTSA, January, 1996 DOT HS 808 338.

The linkage creates a permanent linked record which can be used now or in the future. It expands the usefulness of each individual state data file beyond its original data set to include all of the computerized data elements in each of the files being linked. When sanitized, the linked data are and can be used by multiple users, including highway safety, the medical community, injury control and others, for multiple purposes. The more the linked data are used, the more state data quality improves, definitions are standardized, and collaboration is strengthened.

What Resources Do States Need to Implement CODES?

Existing state data resources can be used for data linkage. The cost of the linkage process depends on the availability and quality of the state data. As a minimum, the following resources are required.

- Computerized, population-based, statewide crash and injury data
- Sufficient identifiers to discriminate among the crashes and the occupants involved in the crash.
- Workstation or micro-computer, depending on the volume of records being linked.
- Staff with expertise related to the state data files (crash, EMS, hospital), computer hardware and data file management.
- Personnel time equivalent to about 2 full time equivalents for a minimum of one month if data are available and edited. Additional time will be required if data are missing or require editing.



What Resources Are Available to Assist States Interested in Implementing CODES?

Your NHTSA Regional Administrator will provide upon request access to experts from the CODES states for customized technical assistance on-site or via telephone. This assistance covers information on how to obtain state data, resolve barriers related to confidentiality and privacy, prepare the files for linkage, and establish an advisory committee to institutionalize data linkage. Experts also are available to help your state implement the probabilistic linkage software, perform the linkage, validate the linkage results and assist in developing analytical uses for your linked state data.

REGION 1—Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont
George A. Luciano
617-494-3427

REGION 2—New Jersey, New York, Puerto Rico, Virgin Islands
Thomas M. Louizou
914-682-6162

REGION 3—Delaware, District of Columbia, Maryland, Pennsylvania, Virginia, West Virginia
Eugene Peterson
410-962-1774

REGION 4—Alabama, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, Tennessee
Thomas J. Enright
404-562-3739

REGION 5—Illinois, Indiana, Michigan, Minnesota, Ohio, Wisconsin
Donald J. McNamara
708-503-8822

REGION 6—Arkansas, Louisiana, New Mexico, Oklahoma, Texas, and Indian Nations
Georgia S. Chakiris
817-978-3653

REGION 7—Iowa, Kansas, Missouri, Nebraska
Troy R. Ayers
816-822-7233

REGION 8—Colorado, Montana, North Dakota, South Dakota, Utah, Wyoming
Louis R. DeCarolis
303-969-6917

REGION 9—American Samoa, Arizona, California, Guam, Hawaii, Nevada, Mariana Islands
Joseph Cindrich
415-744-3089

REGION 10—Alaska, Idaho, Oregon, Washington
Curtis A. Winston
206-220-7640

- Reports Distributed by NHTSA

So You Want to Link Your State Data—
Provides information about data access and file preparation for linkage.

The Report to Congress on the Benefits of Safety Belts and Motorcycle Helmets—
Presents results of an analysis using linked data from the seven CODES states.

CODES Research Note—
Describes the statistical model used by CODES to generate the results reported in the Report to Congress on the benefits of safety belts and helmets.

CODES Technical Report—
Describes the linkage and state specific analysis performed for the CODES Report to Congress.

These reports may be requested by calling: Dennis Utter, CODES COTR, 202-366-5351 Sandy Johnson, CODES Consultant, 202-366-5364—NHTSA, 400 Seventh St., SW, Room 6125, Washington, DC 20590, or FAXing your request to 202-366-7078

The Report to Congress can be downloaded from the World Wide Web at: http://www.nhtsa.dot.gov/people/ncsa/reports.html

- Report Distributed by the National Association of Governor's Highway Safety Representatives

Final Report: Strategic Planning Meeting on CODES and Data Linkage, August 4-5, 1995
Report may be obtained from:
Barbara Harsha, Executive Director
National Association of Governors' Highway Safety Representatives
750 First Street, NE, Suite 720
Washington, DC 20002
202-789-0942